

An assessment of the difficulty of questions used in the ISSP-questionnaires, the clarity of their wording, and the comparability of the responses

Zouwen, Johannes van der

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Zouwen, J. v. d. (2000). An assessment of the difficulty of questions used in the ISSP-questionnaires, the clarity of their wording, and the comparability of the responses. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 46, 96-114. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-199369>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

An assessment of the difficulty of questions used in the ISSP-questionnaires, the clarity of their wording, and the comparability of the responses

by Johannes van der Zouwen¹

Zusammenfassung

International vergleichende Umfrageforschung basiert auf der Annahme, daß erstens die Antworten der Befragten in ausreichendem Maße Auskunft geben über das Verhalten und die Meinungen, die im Fragebogen abgefragt werden, und zweitens daß die Antworten der Befragten über die verschiedenen Länder hinweg ausreichend vergleichbar sind, trotz der Unterschiede in den verwendeten Sprachen. Im Rahmen dieser Arbeit wird untersucht, ob diese Annahme für eine große Auswahl von Umfragen gerechtfertigt ist, die im Kontext des International Social Survey Program (ISSP) entstanden sind. Auf der Basis eines konzeptionellen Modells, basierend auf methodischer Forschung zu Responseeffekten, sollen die Fragebögen bewertet werden. Es kann gezeigt werden, daß die Qualität der Umfrage von der Schwierigkeit und Klarheit der Fragen ebenso abhängig ist wie von der Länge des Fragebogens sowie der Präsenz oder Abwesenheit von verzerrenden Faktoren. Die Fragen des ISSP werden anhand von zehn Kriterien bewertet, die den Schwierigkeitsgrad der Fragen widerspiegeln und anhand von drei Indikatoren für die Eindeutigkeit der Fragestellungen. Die Vergleichbarkeit der Antworten wurde unter verschiedenen Aspekten bewertet, u.a. anhand der „Vergleichbarkeit der Worte in der Frageformulierung“ und anhand der „Ähnlichkeit der zu bewertenden Objekte“. Die Ergebnisse zeigen, daß einige Fragen in ISSP-Surveys ziemlich schwierig und andere unklar formuliert sind. Darüber hinaus ist die Vergleichbarkeit der Antworten reduziert, da z.T. unterschiedliche Intervieweranweisungen in den teilnehmenden Ländern gegeben wurden, und weil einige Fragen zu den Einstellungen in den Ländern unterschiedliche Bedeutungen haben.

¹ **Johannes van der Zouwen** is professor of social research methodology at the Faculty of Social Cultural Sciences of the Vrije Universiteit, De Boelelaan 1081c, 1081 HV Amsterdam, The Netherlands (email: zouwen@scw.vu.nl). This article is a revised version of a paper presented at the International Conference on Large Scale Data Analysis (LSDA-conference) at the Zentralarchiv für Empirische Sozialforschung, University of Cologne, May 25-28, 1999. The ISSP-data utilised in this paper were documented and made available (via the CD-ROM "ISSP 1985-1995 Data and Documentation") by the Zentralarchiv für Empirische Sozialforschung. The data for the 'ISSP' were collected by independent institutions in each country. Neither the original collectors nor the Zentralarchiv bear any responsibility for the analyses or interpretation presented here.

Abstract

The rationale of cross-national survey research rests on the assumptions (1) that the answers given by respondents are sufficiently informative about the behaviours and opinions asked for in the questionnaires, and (2) that the answers of respondents participating in surveys, conducted in different countries, are sufficiently comparable, despite differences in the languages used. It is investigated whether these assumptions are justified for a large selection of surveys conducted within the context of the International Social Survey Program (ISSP). A conceptual model based on methodological research after response effects guides this evaluation of questionnaires. The model asserts that response quality depends on the difficulty and clarity of the questions, on the length of the questionnaire and on the presence or absence of biasing factors. Questions from ISSP-questionnaires are evaluated by using ten criteria relating to question difficulty and three indicators concerning the clarity of the questions. The comparability of responses to these questions is evaluated by looking at six different aspects, like 'equivalent wording of the question' and 'similarity of the objects to be evaluated'. This evaluation shows that some questions asked in the ISSP-surveys are rather difficult while others are not very clear. The comparability of the responses is also hampered by the fact that different modes of questionnaire administration are used in the participating countries and that the attitude objects to be evaluated by respondents may differ between questionnaires.

1 Introduction

Ten years ago Øyen (1990, p. 1) observed that “more cross-national studies than ever before are being carried out, and the need as well as demand for comparisons across countries is formidable”. Seven years later, **Johnson** et al. (1997, p. 87), start their paper by saying: “Recent years have seen a significant expansion in the variety and number of surveys conducted across national and cultural boundaries.” International comparative survey research has become a ‘booming business’.

The larger the number of countries participating in cross-national research programs, and the more these countries vary with respect to the ways the separate surveys are conducted, the more urgent becomes the issue of the *comparability* of the responses to questions from different questionnaires. And even prior to the issue of the comparability of the responses we have the issue of the *quality* of the responses, that is, the degree to which the responses given are informative about the states, behaviours or opinions asked for. The answer of a respondent is informative if the answer is substantive and complete, and if the answer is hardly unaffected by response effects or by response bias.

In other words, the rationale of international comparative research projects, in which data are compared, collected in different countries by means of questionnaires containing (at least partly) the same questions, rests on at least the following assumptions:

1. Answers given by respondents are sufficiently *informative* about the states, behaviours and opinions asked for in the relevant parts of the questionnaires, and
2. Answers are sufficiently *comparable* between the questionnaires, despite differences in the languages used.

In the present paper I will try to investigate in how far these two assumptions are justifiable, by analysing the content, format and structure of the questionnaires used in a current comparative research programme, i.e., the International Social Survey Programme (ISSP).

The ISSP is a continuing, annual program of cross-national collaboration between national research groups. Each of the over twenty countries participating in this program undertakes to include common questions as a supplement to their regular national surveys. The topics change from year to year by agreement, and are replicated every five year or so. Some examples of topics are the role of government, social inequality, family and changing sex roles, religion, and environment (*Uher*, 1998, 1999).

For this paper I have analysed those questionnaires that were written in Dutch, English, French and German.² The 77 selected questionnaires form about half (49%) of the total number of 157 questionnaires.

2 A conceptual model for evaluating the quality of questions

In order to evaluate the quality of the questions used in the ISSP questionnaires, a conceptual model³ is used (*Van der Zouwen* and *Dijkstra*, forthcoming). This model is based on research on response effects of question content (*DeLamater*, 1982), research on question comprehension by *Belson* (1982) and research on cognitive processes in question answering, by *Schwarz* and *Sudman* (1996), *Sudman*, *Bradburn* and *Schwarz* (1996), and *Tanur* (1991). The model is also based on methodological rules for question wording, as formulated by *Sudman* and *Bradburn* (1982) and by *Foddy* (1993).

In this conceptual model the dependent variable is RESPONSE QUALITY. The main determinants of response quality are the DIFFICULTY of the question, the CLARITY of the task involved, and the presence or absence of response BIASING FACTORS.

2 The reason for this selection is purely practical: these four languages are the only ones the author can read.

3 By 'conceptual model' I mean a type of model that functions as a heuristic and as a means to structure the evaluation of the ISSP-questionnaires. An empirical test of this model is not intended in this paper.

The **DIFFICULTY** of the question and the **CLARITY** of the task affect the occurrence of random errors and thus the reliability of the responses, while the presence of **BIASING FACTORS** may lead to systematic errors.

Intervening and additional variables in this conceptual model are:

The **PROBABILITY** of the occurrence of mistakes and misunderstandings made by the respondent,

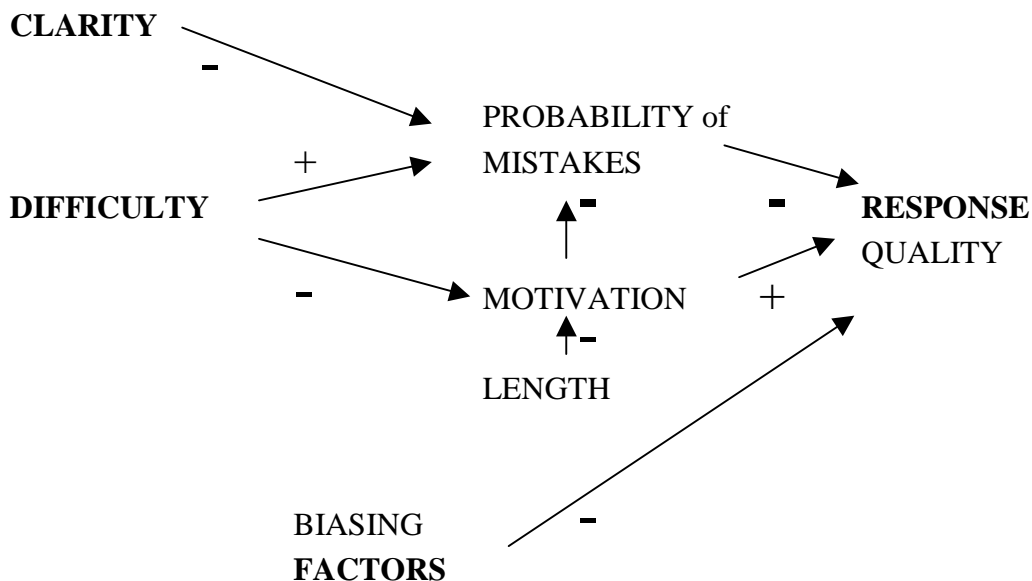
the **MOTIVATION** to properly fulfil the task as a respondent, and

the **LENGTH** of (parts of) the questionnaire.

The **PROBABILITY** of mistakes is negatively affected by the **CLARITY** of the question and positively affected by the **DIFFICULTY** of the question. The **MOTIVATION** of the respondent is negatively affected by the **DIFFICULTY** of the question and by the **LENGTH** of the questionnaire.

The assumed relations between these variables are depicted in Figure 1.

Figure 1. Variables affecting **RESPONSE QUALITY** in surveys



3 The **DIFFICULTY** of the question

According to our conceptual model, the difficulty of a question has a negative effect on the quality of the response to that question. Methodological research using cognitive interviewing (e.g. *Sudman, Bradburn and Schwarz, 1996*) and analyses of the interactions

between respondent and interviewer (cf. *Van der Zouwen* and *Dijkstra*, 1995, 1996a, 1996b, forthcoming) have shown that the difficulty of a question will (*ceteris paribus*) be greater

1. If more cognitive operations are necessary to 'compute' the response.
2. If it is necessary to retrieve information from long term memory.
3. If too much information about the question and the response alternatives has to be stored in short-term memory in order to adequately answer the question.
4. If the question relates to a 'hypothetical', or a future, situation.
5. (For open questions) the more precision of the response is required.
6. If (for closed questions) the response alternatives offered don't fit well with the question.
7. (For closed questions) if the response alternatives are unevenly distributed over the range of possible values of the variable that the question is meant to measure.
8. If no use is made of 'responding aid devices' (like show cards, visual aids, 'ladder/barometer').
9. If the topic of the question involves problems of self-presentation (topic threat or large differences with respect to the social desirability of response alternatives).
10. If the information asked for is hardly accessible.

In the next part of this section these ten indicators of question difficulty will be applied to the questions included in the ISSP-questionnaires used from 1985 through 1995. Most examples are drawn from the British versions of the ISSP-questionnaires. Questions are indicated by a form of short hand: GB-86/9 means question 9 from the survey conducted in 1986 in Great Britain (ZA-Studien-Nr. 1620).

3.1 DIFFICULTY-1: Responding requires many cognitive operations.

The ISSP-questionnaires contain some questions that require a lot of cognitive 'computation'. Examples are

- a. "*Thinking now of close friends – not your husband, or wife, or partner, or family members – but people you feel fairly close to. (a) How many close friends would you say you have? (b) How many of these friends are people you work with now? (c) How many of these friends are your close neighbours?*" [GB-86/9a-c]. In order to answer questions (b) and (c) properly, the respondent first has to 'observe' his social network, and look for those people (s)he feels 'fairly close to'. Next, he or she has to subtract from this set the partner and family members. Then select from the remaining people those with whom respondent is working with presently, respectively are close neighbours, and finally report the number of people belonging to these subsets. This procedure for the estimation of the size of part

of the personal network is not only difficult for the respondent but it also may lead to underreporting (*Broese van Groenou* and *Van Tilburg*, 1996).

Even more difficult is the following question:

b. *“Please think of your present job (or your last one if you don’t have one now). If you compare this job with the job your father had when you were 16, would you say that the level or status of your job is (or was): Much higher than your father’s/ Higher/ About equal/ Lower/ Much lower than your father’s/ (I never had a job)/ (Never knew my father/father never had a job)”*. [GB-87-12 & GB-92/11].

In order to answer that question properly, respondents first have to retrieve from long term memory information about their father’s job when the respondents were sixteen and assess the status of that job by criteria used at that time (or by criteria used nowadays?). Next they have to assess their own jobs, and compare both assessments. Eventually they have to express the estimation of this difference by choosing one of the five response categories offered. The probability that somewhere in this cognitive processing a mistake is made which affects the quality of the response is quite large.

3.2 DIFFICULTY-2: Responding requires retrieval from long term memory (retrospective questions).

Retrospective questions are 'notorious' for the occurrence of memory errors that negatively affect response quality (*Van der Vaart*, *Van der Zouwen* and *Dijkstra*, 1995). In ISSP-questionnaires related to the social mobility and to the religious background of the respondent, a lot of retrospective questions have been asked. For example:

a. *“Did your mother ever work for pay as long as one year after you were born and before you were 14? Yes, she worked/ No/ Did not live with mother”* [GB-88/15 & GB-94/C2.08].

b. *“What was your mother’s religion, if any, when you were a child? No religion/ Christian - no denomination/ Roman Catholic/ Church of England/Anglican” plus 15 other response categories*. [GB-91/A2.22].

c. *“When you were a child, how often did your mother attend religious services?: Never/ Less than once a year/ About once or twice a year/ Several times a year/ About once a month/ 2-3 times a month/ Nearly every week/ Every week/ Several times a week/ Never knew mother/does not apply/ Can’t say/Can’t remember”* [GB-91/A2.25].

The frequency of mother’s church visits and the precise church affiliation may be not very salient topics for a child while the recall period may be very long: thirty years or even more. The combination in one question of low saliency of the topic with a long recall period may

threat the quality of the responses to that question, the more so because no recall aiding devices are used (*Van der Vaart*, 1996)

3.3 DIFFICULTY-3: Adequate responding requires storing too much information about the question in short-term memory.

In order to answer a question adequately, respondents have to have a complete overview in their 'short-term memory' of the text of the question and the corresponding response alternatives. Otherwise primacy or recency effects may occur; that is, respondents then react mainly to the first or last part of the question (cf. *Foddy*, 1993). ISSP-questions that by far exceed the capacity of short-term memory are for example:

- a. *"Some people think the government should provide financial assistance to university students. Others think the government should not provide such aid. In each of the circumstances listed below should the government provide grants that would not have to be paid back, provide loans which the student would have to pay back, or should the government not provide any financial assistance? For students whose parents have a low income: Government should give grants/ Government should make loans/ No Government assistance/ Can't choose". [GB-85/17].* Before respondents can give any answer they have to read (or, in some countries, listen to) over eighty words.
- b. *"Suppose the police get an anonymous tip that a man with a long criminal record is planning to break into a warehouse. Do you think the police should be allowed, without a Court Order, to keep the man under surveillance? Definitely/ Probably/ Probably not/ Definitely not/ Can't choose". Likewise: 'to tap his telephone, 'to open his mail', and 'to detain the man overnight for questioning'. [GB-90/20.5].* For this last question one may hope that most respondents have noticed the limiting condition "without Court Order", otherwise they respond in fact to a much more extreme item.

3.4 DIFFICULTY-4: The question relates to a 'hypothetical' or a 'future' situation'.

Methodological research of interactions between interviewers and respondents in survey-interviews (e.g., *Van der Zouwen* and *Dijkstra*, 1996a) has shown that a lot of respondents have problems with answering hypothetical questions ("What would you do if you were X?" ... " But I am not X!") or with questions about future situations. However, quite a lot of ISSP-questions refer to those hypothetical or future situations. To mention only a few examples:

- a. *"If incomes became more equal in Britain, some people would get higher incomes and some would get lower incomes. Do you think your income: would definitely go up/ would probably go up/ ..."* [GB-92/14].

b. *“Suppose you were unemployed and couldn’t find a job. Which of the following problems do you think would be the worst? Lack of contact with people at work? Not enough money? Loss of self-confidence? ...”* [GB-89/7].

c. *“Within the next ten years, how likely do you think it is that there will be a large increase in ill-health in Britain’s cities as a result of air pollution caused by cars? Certain to happen/ Very likely to happen/ Fairly likely to happen/ Not very likely to happen/ or - Certain not to happen”* [GB-93/2.12c].

The problem with these particular questions is that respondents may have difficulty with accepting the reality of the hypothetical situation (‘incomes will not become more equal’; ‘I am not unemployed’) or have difficulty with imagining a situation in a far remote future (‘the next ten years’).

3.5 DIFFICULTY-5: The more precision of the response is required (for open questions).

The ISSP-questionnaires contain only a few open questions. But responding to these few open questions requires a lot of precision. For example:

“We would like to know what you think people in these jobs actually earn. Please write in how much you think they usually earn each year, before taxes. (Many people are not exactly sure about this, but your best guess will be close enough. This may be difficult, but it is important, so please try.)” [GB-87/5 & GB-92/4]. Jobs for which an estimate is asked are ‘a bricklayer’, ‘a doctor in general practice’, ‘a bank clerk’, ‘the owner of a small shop’, ‘the chairman of a large national company’, ‘a skilled worker in a factory’, ... ‘a cabinet minister in the national government’, ... and (in 1992) ‘an Appeal Court judge’. The required precision actually depends upon the local currency, but is even for Great Britain (pound sterling!) far greater than the vague ideas people have about earnings in hardly unknown jobs (like an Appeal Court judge). Even if the researcher gets answers to these questions, the informative content of these answers will be modest.

3.6/7 DIFFICULTY-6&7 Inadequacies regarding (the set of) response alternatives.

For the large majority of the ISSP-questions the response categories fit quite well with the preceding texts of the questions. Only in a very few occasions one may observe some inadequacies regarding the text of the response alternatives. For example:

“If the government had to choose between keeping down inflation or keeping down unemployment, to which do you think it should give the highest priority?”

Keeping down inflation? / Keeping down unemployment? / Can't choose." [GB-90/2.12].

Both response alternatives are, 'logically speaking', not mutually exclusive. The respondent may wish that a government kept down unemployment by keeping down inflation. These respondents are forced to make an unrealistic choice. But again, on the evaluation criterion 'adequacy of the response alternatives' the ISSP-questionnaires score quite well in general.

3.8 DIFFICULTY-8: No use is made of 'responding aid devices' (show card, visual aids, ladder/barometers).

According to the official documentation of ISSP (see Table 1), the questionnaires are administered either:

- in the form of a self-completion questionnaire (in 4, 6 respectively 10 participating countries in respectively 1985, 1990 and 1995),
- during a face-to-face interview (in 2, 4 and 14 participating countries in respectively 1985, 1990 and 1995), or
- in either way (Italy in 1985 through 1990).

Table 1 Mode of administration (S = self-completion or F = face to face interview) and response rate for ISSP-surveys* conducted in 1985, 1990 and 1995 respectively.

Response Rate	1985		1990		1995		total	
	S	F	S	F	S	F	S	F
40 <50%	D						1	0
50 <60%			D-W GB		D		3	0
60 <70%	AUS GB		D-E N		GB N, S	RUS, A CZ, NL	7	4
70 <80%	USA	A		IRL	CDE CDF NZ USA	J LV	5	4
80 <90%			AUS USA		PL	BG SLO	3	2
90 <100%				H		SK	0	2
Unknown		I**		I* IL	AUS	H, IRL I, RP, E	1	8
Total	4	2	6	4	10	14	20	20

*Countries are indicated by the international automobile identification codes.

** : In Italy the questionnaire is either administered in an interview or self-completed

In case it is officially indicated that a face-to-face interview is used:

- this information is often incorrect: actually not a face-to-face interview, but a self-completion questionnaire, is used to administer the ISSP-questionnaire (e.g., in The Netherlands), or
- the questionnaire is reformatted into a proper face-to face-questionnaire with show cards (Austria), or
- the questionnaire is not adequate: questionnaires to be filled out by the respondents themselves require a type of formatting which greatly differs from questionnaires used by interviewers in face-to-face interviews. Many of the questions are too long and/or complex to be used in a 'normal' face-to-face interview without making extensive use of show cards (see also subsection 3.3 above).

3.9 DIFFICULTY-9: Problems of self-presentation for the respondent.

In their classical study of response effects in surveys, *Sudman* and *Bradburn* (1974) have shown that sometimes the respondent faces problems of self-presentation: an honest answer may be a socially undesirable one at the same time. This component of task difficulty is hardly existent in the ISSP-surveys, provided the questionnaire is self-completed by the respondent! If the questionnaire is used in an interview, social desirability may sometimes play a role (e.g., with questions concerning 'environmental behaviour' there may be some bias in the direction of 'good' behaviour).

3.10 DIFFICULTY-10: The information asked for is hardly accessible

Some of the ISSP questions have the character of 'knowledge questions' or 'quiz items'. For example:

- a. *"For each statement below, just tick the box that comes closest to your opinion of how true it is: definitely true/ probably true/ probably not true/ definitely not true/ can't choose"*.

Statements are e.g.: 'All radioactivity is made by humans' (2.10a); 'Antibiotics can kill bacteria but not viruses' (2.10b); 'The greenhouse effect is caused by a hole in the earth's atmosphere' (2.11c) [GB-93/2.10 & 11]. Or,

- b. *"In general, do you think that a rise in the world's temperature caused by the 'greenhouse effect' is: extremely dangerous for the environment/ very dangerous/ somewhat dangerous/ not very dangerous/ or, not dangerous at all for the environment?/ Can't choose"* [GB-93/17a].

For many respondents it will be hardly possible to give the ‘good’ answer to these questions. This may hamper the motivation to answer subsequent questions: “Now I know for sure that I have not enough knowledge to participate in this survey”.

Another example of questions requiring information the ‘modal’ respondent does not have access to, are the questions regarding estimations of income earned in other jobs (mentioned in section 3.5).

4 The CLARITY of the task for the respondent

As stated in the conceptual model, the clarity of the task for the respondent has a positive effect on the response quality. The clarity of the task will decrease

1. if the question is ambiguous;
2. if the question uses unfamiliar words;
3. if clear indications about proper performance are missing, like instructions regarding turn taking (in face-to-face interviews), routing (the applicability of a particular question to a particular respondent), and the ways of indicating the choice of a response alternative (in self-administered questionnaires).

If we evaluate the ISSP questionnaires with these three aspects of the criterion of CLARITY we come to the following results.

4.1 CLARITY-1: Ambiguity of the question.

a. An example of unclear meaning of wording is the following:

“Do you think the number of immigrants to Britain nowadays should be: increased a lot/ increased a little/ remain the same as it is/ reduced a little/ reduced a lot?/ Can’t choose” [GB -95/11]. In the Dutch version of the questionnaire this question reads: “Denkt u dat het aantal immigranten in Nederland vandaag de dag: sterk moet worden verhoogd/ een beetje moet worden verhoogd/ hetzelfde moet blijven zoals het nu is/ een beetje moet worden verminderd/ sterk moet worden verminderd/ geen mening” [NL-95/13]. The number of ‘immigrants to’ is of course much smaller than the number of ‘immigrants in’, and reducing the number of ‘immigrants in’ requires more drastic measures than the reduction of the yearly influx of ‘immigrants to’. The fact that the response distribution for the Netherlands is about the same as that for Great Britain indicates that some respondents, in both countries, may have misunderstood the question as presented in ‘their’ questionnaire; a clear example of ‘question ambiguity’.

b. ‘double barrelled’ questions or ‘hidden screening’ questions.

In methodological literature (e.g., *Fowler* and *Mangione*, 1990, *Foddy*, 1993), survey researchers are warned for asking two questions (A and B) in one item because if the respondent agrees with A but disagrees with B, he or she will have problems with responding. A similar type of problem arises if in a question a statement A is followed by a reason B for A. If the respondent agrees with A but disagrees with B, a negative answer is not interpretable. Examples of questions in which a statement is followed by a reason for that statement are:

“Inequality continues because it benefits the rich and powerful” [GB-87/4c & GB-92/3c].

“Human beings should respect nature because it was created by God” [GB-93/2.05d].

“There will always be conflict between management and workers because they are really on opposite sides: Strongly agree/ Agree/ Neither agree nor disagree/ Disagree/ Strongly disagree/ Can’t choose” [GB-89/5a].

4.2 CLARITY-2: Unfamiliar words used in the question.

The ISSP-questionnaires seldom use unfamiliar words. But for at least Dutch respondents the definition given of ‘taxes’ (i.e., including national insurance) as in the example below will be quite unfamiliar and may lead to response errors.

a. *“Generally, how would you describe taxes in Britain today (We mean all taxes together, including national insurance, income tax, VAT and all the rest.)” [GB-92/7].*

b. *Sometimes the concepts themselves are rather vague or unfamiliar. I wonder what respondents may have thought by hearing the word ‘this’ in the following question:*

“Sometimes at work people find themselves the object of sexual advances, propositions, or unwanted sexual discussions from co-workers or supervisors. The advances sometimes involve physical contact and sometimes just involve sexual conversations. Has this ever happened to you? Yes/ No/ Never have worked.” [GB-94/C2.16].

c. From the official instructions given to designers of the national questionnaires it becomes apparent that the concept ‘your neighbourhood’ has a different meaning for people living in a small village or living in a big city. Respondents living in small towns have to decide by themselves how they will interpret the word ‘neighbourhood’:

“To begin, we have some questions about where you live: your neighbourhood or village, your town or city, your county, and so on. (By “neighbourhood” we mean the part of the town/city you live in. If you live in a village, we take this as your “neighbourhood”.) How

close do you feel to (a) your neighbourhood (or village): very close/ close/ not very close/ not close at all/ can't choose; (b) your town or city", etc. [GB-95/1].

4.3 CLARITY-3: Absence of clear instructions

The clarity of the task of the respondent decreases if clear instructions about proper administration of the questionnaire are missing. These instructions may concern

- the routing through the questionnaire (and the applicability of a particular question), and
- the interpretation of the questions and the indication of the choice of the response alternative.

Especially with self-administered questionnaires, where help and correction by interviewers is lacking, the layout of the questionnaire is crucial. **Jenkins** and **Dillman** (1997) have formulated the following principles for 'navigational guides used in self-administered questionnaires':

1. Use the visual elements of brightness, colour, shape, and location in a consistent manner to define the desired navigational path for respondents to follow when answering the questionnaire (p. 177).
2. When established format conventions are changed in the midst of a questionnaire, prominent visual guides should be used to redirect respondents (p. 182).

With respect to the first 'principle', no precise overall assessment of the ISSP questionnaires can be given because the layout of the questionnaires does differ quite considerably between participating countries and over time. Anyhow, the layout of the questionnaires becomes more and more 'professional'. Regarding the second 'principle', it can be observed that, mainly for reasons of space, format conventions are changed within a particular questionnaire without any 'warning' to the respondent. For example: in the questionnaire GB-87, sets of 'continuous' response alternatives like 'strongly agree', ... 'strongly disagree' are subsequently ordered: vertically (Q1 - Q3), horizontally (Q4, Q7), vertically (Q8, Q9), horizontally (Q10), vertically (Q11 + Q12), without any visual guides 'to redirect respondents'.

5 The absence or presence of BIASING FACTORS

In general, the questions in the ISSP-questionnaires are formulated in a balanced way. Only in a very few cases one might observe the occurrence of a leading question, in which the respondents are 'warned for' choosing the response alternative 'much more', while other alternatives lack an indication of their consequences. An example is:

“Listed below are various areas of government spending. Please show whether you would like to see more or less government spending in each area. Remember that if you say “much more”, it might require a tax increase to pay for it.

The environment: Spend much more/ Spend more/ Spend the same as now/ Spend less/ Spend much less/ Can’t choose” [GB-85/22 & GB-90/2.11]

Another type of question wording which might bias the responses is the inclusion in the question itself of an assumption for which no evidence is given (‘truth asserting questions’). An example is:

“How proud are you of Britain in each of the following? (..) its fair and equal treatment of all groups in society? (..): Very proud/ Somewhat proud/ Not very proud/ Not proud at all/ Can’t choose” [GB-95/6j].

This question assumes that in Britain (and in each of the other countries participating in ISSP) all groups are treated fair and equal; an assumption that for some respondents may sound too optimistic. Those respondents have difficulty in expressing their opinion, which may lead to response errors.

6 Inadequacies of the questionnaire in its entirety

When evaluating a questionnaire one not only has to look at the separate questions, but also to the questionnaire in its entirety: are topics presented in a ‘logical’ order, and does the part of the questionnaire devoted to a particular issue remain within reasonable limits of time and space? That sometimes the order of the questions leads to a strange sequence of topics is illustrated by questionnaire GB-1993:

The question: *“Please tick one box to show which statement comes closest to expressing what you believe about God” [GB-93/2.22]* is preceded by a series of questions about environmental behaviour, and directly followed by *“Would you describe the place where you live as: a big city/the suburbs or outskirts of a big city/ a small city or town/ a country village/ or, a farm or home in the country”*.

With respect to the length of parts of the questionnaire it struck me that in GB-87/1 the respondent has to evaluate thirteen aspects with regard to their ‘importance for getting ahead in life’; a quite exhausting task, reducing respondent’s motivation from the onset.

In GB-86/1-6 respondents have to answer three detailed questions about contacts with respectively: mother, father, sisters, brothers, daughters, sons, other adult relative with whom R has most contact with, and closest friend. Again, the researcher may need this

information, but it places a heavy burden on the respondent, with the risk of producing errors and ‘missings’.

7 An assessment of the COMPARABILITY of the responses

The second assumption behind cross-national survey research, mentioned in section 1, is that “answers are sufficiently comparable between questionnaires, despite differences in languages used”. I will mention below some aspects of the COMPARABILITY of the questionnaires, and therewith of the data collected by different versions of the questionnaire.

The investigation of ‘subjective phenomena’ requires some sort of black box analysis, because opinions and attitudes are not directly observable. By asking questions (input) and listening to, or reading, the answers (output) one may, under certain assumptions, make valid inferences about the respondents’ opinions and attitudes (the content of the black box). One of the assumptions behind the survey methodology is that when a group of respondents answer the same question, those who give the same answer have the same value on the unobservable variable and those who give different answers have different values. This logic of reasoning requires that the researcher offers the same stimulus (question) to all respondents; for if the stimulus varies, variation in responses may be due to the variation in the stimulus rather than to the variation in the unobserved variable. In that case the output, i.e., the response, cannot be used for making inferences about the unobservable variable we are trying to measure (*Van der Zouwen*, forthcoming).

7.1 Equivalent WORDING of the question

Thus it is crucial that corresponding questions have an equivalent wording. Otherwise the responses to a particular question in version A cannot be compared in a meaningful way to responses given to the matching question in version B. That this principle is not always followed, can easily be seen by scanning through the code books, where most discrepancies between questionnaires are mentioned. For example: In GB-85/3c the respondent is asked, “*should it be allowed to organise protest marches and demonstrations*”. In the versions for Germany and Austria the statement continues with “*which prevent traffic*”, making it more ‘easy’ for these latter respondents to answer with ‘no’; by the same token making the responses of the German and Austrian respondents incomparable with the respondents from other countries.

Another example:

“There are some people whose views are considered extreme by the majority. First consider people who want to overthrow the government by revolution. Do you think such peo-

ple should be allowed to teach 15 years olds in school?" [GB-85/14a_{ii}]. However, in the Italian questionnaire the children involved are not 15 years old but 18 years old, making the statement more easily to affirm.

Still another example of differences between question wording was mentioned in section 4.1 where it was observed that the Dutch questionnaire asks for 'immigrants in' while the British questionnaire deals with 'immigrants to'.

7.2 Similar WORDING and CODING of response categories for 'can't choose', 'don't know', 'no answer'.

The code books of the ISSP-surveys show that in different countries different wordings and codes are used for non-substantive response categories like 'can't choose', 'don't know', 'no opinion' etc. For example: In NL-95 'can't choose' changed into 'geen mening' (no opinion) and in NL-98 changed into 'weet niet' (don't know).

7.3 Same 'CONTEXT' of the question.

In methodological research it has been established that preceding questions may have impact on the responses to the present question (*Schwarz* and *Sudman*, 1992). To avoid these context effects, or order effects, it is important that the respondents in different countries answer the same series of questions. Due to the fact that many countries skip parts of the ISSP-questionnaire, or add some other ones to the questionnaire, respondents in different countries have responded to another set of preceding questions, therewith hampering the comparability of the responses to the present question.

7.4 Equal MODE of administration of the questionnaire.

Research after 'mode effects' has shown that the mode of administration has an effect on the responses obtained (*De Leeuw*, 1992). Apart from the issue whether a particular questionnaire can be used adequately in face-to-face interviews, it is likely that responses to questions in self-administered questionnaires are slightly different from those given in face-to-face interview, thus hampering their comparability. From Table 1 above it is clear that the problem of different modes of administration used for the ISSP-survey has not been solved as yet. It seems that the growth of the number of participating countries has led to an increase of face to face interviewing, therewith aggravating the problem that many questions from the ISSP-questionnaires are too difficult to be answered without using response aiding devices like show cards.

7.5 Similarity of the 'OBJECTS' to be evaluated.

In many questions the respondent have to evaluate a certain '(attitude) object'. If these objects differ for different respondents, the evaluations cannot be compared meaningfully. Sometimes this incomparability is quite obvious, e.g., in those questions in which the 'object' to be evaluated is 'country specific'. For example: "In Britain what you achieve in life depends largely on your family background: Agree strongly/ Agree/ Neither agree nor disagree/ Disagree/ Disagree strongly" [GB-85/10C]. If one changes 'Britain' for the name of another country, i.e., Germany, the question itself has changed. It looks as if the roles of 'respondent' (speaking for himself) and 'informant' (speaking about the group one belongs to) are mixed up in these questions (*Back* and *Cross*, 1982).

A more complicated situation occurs if agreement or disagreement with a statement concerning an attitude object (e.g., one's own country), is the result of both an intrinsic characteristics of that object (e.g., the military power of that country) and of the value on some latent trait within respondents (their chauvinism). In that case the answers, given to questions with different attitude objects, are really incomparable. The situation resembles one equation with two unknowns. To illustrate this point one may look at the responses given to the question "How proud are you of [one's country] in each of the following" ... "Its political influence in the world" [95/6b]. As can be expected, the percentage respondents from the USA choosing '(very) proud' is high: 80%, in Great Britain somewhat lower (GB = 55%), and quite low for the small country Slovenia (SLO = 29%). But how does one explain the high percentages of New Zealand (NZ = 71%), and Norway (N = 78%)?

7.6 Similarity of the survey-samples

If response distributions for one country, or for one year, are compared with those for another country or another year, the underlying assumption of cross-national and longitudinal research is that the samples from which these distributions stem can be compared meaningfully. This means for example that the sampling procedures are about similar and that the response rates are sufficiently high. Although it is hard to tell whether the sampling procedures used in the different ISSP countries are sufficiently alike, it is easy to see (in Table 1) that the response rates show a bewildering variation: from below fifty percent (Western Germany in 1985) to over 90% in Hungary⁴ and in the Slovakian Republic, with most surveys in between 60 and 80%.

4 It is tempting to try to explain the enormous variety in response rates between participating countries, and especially the very high response rates reported for Hungary and the Slovakian Republic. However, the documentation of these surveys in the Codebook of ZA Study 2880 is not detailed enough to construct a valid explanation.

8 An over-all assessment of the ISSP-questionnaires

This paper was intended to test two basic assumptions behind cross-national research, i.e. that the responses obtained in ISSP-surveys conducted in different countries are informative as well as comparable. The results of the evaluation of the ISSP questionnaires can be summarised by the following statements:

1. Only a few questions are incorrectly formulated, i.e., are leading questions, or double barrelled questions, include unjustified assumptions, or have an incomplete set of response alternatives.
2. However, quite a lot of questions involve long and complex texts and are hardly fit for use in face-to-face interviews, and are problematic even in self-completion questionnaires.
3. The information requested from respondents in some questions is hardly accessible to them: retrospective questions about their parents; knowledge questions about impacts on the environment; estimations of the gross income earned in other jobs.
4. As long as the ISSP-questions are answered in the absence of an interviewer no problems of 'self-presentation' are to be expected. However, the data from table 1 show that in a growing number of ISSP-surveys face to face interviews are used, thus increasing the risk of socially desirable responding.
5. The ISSP-questionnaires are generally quite long and working through these series of questions requires a strong motivation of the respondent.
6. Comparisons of responses given to 'country-specific' questions are very risky. Here we have the situation of 'one equation with two unknowns'.

My recommendations for ISSP are: avoid too difficult to answer questions, be careful with the interpretation of the responses to 'country specific' questions, and try to reach a common mode of administration of the questionnaires.

References

- Back, Kurt W. / Cross, Teresa S.** 1982: Response effects of role restricted respondent characteristics, in: **Dijkstra, W.** and **Van der Zouwen, J.** (eds.): *Response Behaviour in the Survey-interview*. London: Academic Press, 189-207.
- Belson, William A.** 1982: *The Design and Understanding of Survey Questions*. Aldershot: Gower.
- Broese van Groenou, Marjolein / Van Tilburg, Theo**, 1996: Network Analysis, in: Birren, J.E. (Ed.) *Encyclopedia of Gerontology*, Volume 2, 197-210. San Diego: Academic Press.
- DeLamater, John**, 1982: Response-effects of question content, in: **Dijkstra, W.** and **Van der Zouwen, J.** (eds.) *Response Behaviour in the Survey-interview*. London: Academic Press, 13-48.
- De Leeuw, Edith D.**, 1992: *Data Quality in Mail, Telephone and Face to Face Surveys*. PhD dissertation Vrije Universiteit, Amsterdam.

- Foddy, William**, 1993: Constructing Questions for Interviews and Questionnaires. Cambridge: Cambridge University Press.
- Fowler Jr., Floyd W. / Mangione, Thomas W.**, 1990: Standardised Survey Interviewing. Newbury Park: Sage.
- Jenkins, Cleo R. / Dillman, Don A.** 1997: Towards a theory of self-administered questionnaire design, in: **Lyberg, Lars** et al. (eds.) Survey Measurement and Process Quality. New York: Wiley, 165-196
- Johnson, Timothy / O'Rourke, Diane / Chavez, Noel / Sudman, Seymour / Warnecke, Richard / Lacey, Loretta / Horm, John** 1997: Social cognition and responses to survey questions among culturally diverse populations, in: **Lyberg, Lars** et al. (eds.) Survey Measurement and Process Quality. New York: Wiley. 87-113.
- Øyen, Else** (Ed.), 1990: Comparative Methodology; Theory and Practice in International Social Research. London: Sage.
- Schwarz, Norbert / Sudman, Seymour** (eds.) 1992: Context Effects in Social and Psychological Research. New York: Springer.
- Schwarz, Norbert / Sudman, Seymour** (eds.) 1996: Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research. San Francisco: Jossey-Bass.
- Sudman, Seymour / Bradburn, Norman M.** 1974: Response Effects in Surveys. Chocago: Aldine.
- Sudman, Seymour / Bradburn, Norman M.**, 1982: Asking Questions; A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass.
- Sudman, Seymour / Bradburn, Norman.M. / Schwarz, Norbert** (eds.)1996: Thinking about Answers; The Application of Cognitive Processes to Survey Methodology. San Francisco: Jossey-Bass.
- Tanur, Judith M.** (ed.) 1991: Questions about Questions; Inquiries into the Cognitive Bases of Surveys. New York: Russell Sage Foundation.
- Uher, Rolf**, 1998: Neues vom ISSP. ZA-Information 42, Mai 1998, 60-66.
- Uher, Rolf**, 1999: Organisational and Methodological Aspects of the International Social Survey Programme (ISSP). Paper presented at the conference on Large Scale Data Analysis (LSDA), University of Cologne, May 25-28, 1999.
- Van der Vaart, Wander**, 1996: Inquiring into the Past: Data Quality of Responses to Retrospective Questions. PhD dissertation Vrije Universiteit Amsterdam.
- Van der Vaart, Wander , van der Zouwen, Johannes / Dijkstra, Wil** (1995): Retrospective questions; data quality, task difficulty, and the use of a checklist. Quality and Quantity 29, 299-315.
- Van der Zouwen, Johannes**, forthcoming: Why study interaction in the survey interview, in: **Maynard, D.W.** et al. (eds.), Standardisation and Tacit Knowledge: Interaction and Practice in the Survey Interview.
- Van der Zouwen, Johannes / Dijkstra, Wil**, 1995: Trivial and Non-trivial Question-Answer Sequences; Types, Determinants and Effects on Data Quality. Proceedings of the International Conference on Survey Measurement and Process Quality in Bristol (UK), pp. 81-86.
- Van der Zouwen, Johannes / Dijkstra, Wil** 1996a: The Impact of the Question on the Interactions in Survey-Interviews. Paper presented at the Fourth International ISA Conference on Social Science Methodology (Essex '96), July 1-5, University of Essex, Colchester UK.
- Van der Zouwen, Johannes / Dijkstra, Wil**, 1996b: Testing (CATI-) Questionnaires Using Computer Assisted Interaction Coding. Paper presented at the International Conference on Computer-Assisted Survey Information Collection (InterCASIC '96), December 11-14, San Antonio, Texas.
- Van der Zouwen, Johannes / Dijkstra, Wil** (forthcoming): Testing questionnaires using interaction coding, in: **Maynard, D. W. et al** (eds.), Standardisation and Tacit Knowledge: Interaction and Practice in the Survey Interview.